Universiteit Utrecht

Nikhef

Instite for Subatomic Physics
Department of Physics and Astronomy
Faculty of Science, Utrecht University

Study of event shape in pp collisions at LHC energies using the PYTHIA model

Bachelor Thesis

Author:

Jim Carstens
Student Number: 5558816
Physics and Astronomy

Supervisor:

Dr. Panos Christakoglou
Nikhef, Utrecht University

16th January 2018

Abstract

Recent experimental results from the analysis of pp collisions at the LHC have sparked much interest since they resemble the ones reported in heavy-ion collisions. The latter are consistent with the creation of a hot and dense coloured matter, the quark-gluon plasma (QGP). These results pose an intriguing question related to their physics origin and force people to look in pp collisions in more detail. In this report, we analyse simulated data of pp collisions at $\sqrt{s} = 7$ TeV for two sets of Monte-Carlo generated collision data using PYTHIA. In particular we focus on different shapes (i.e. jetty vs spherical events) of pp collisions in an attempt to distinguish between phenomena that can be described by QCD or by phenomenological approaches. It is found that high multiplicity events have preferentially high sphericity. Moreover, high p_T events are jetty, and low p_T events are spherical. In one of the two datasets colour reconnection (CR) is turned on, and in the other off. This mechanism is one of the phenomenological principles that can explain the QGP-like behaviour in pp collisions. Although only a small difference is found, CR could explain a rise in p_T over multiplicity.

Contents

1 Introduction

One of the most important predictions from the Standard Model of particle physics is the occurrence of certain phase transitions in matter at extreme conditions. From these transitions, the one predicted by the theory of quantum chromodynamics (QCD) is possible to be produced and studied in the laboratory. This is the transition from the matter we know in our daily world to a quark-gluon plasma (QGP), which is also what the Universe went through during a crucial period of the Big Bang. It can be probed in ultra-relativistic collisions of nuclei in particle accelerators. However, since this phase of matter quickly dissociates and decays, indirect observation methods have to be used, relying on theory from various physics sub-fields and advanced simulation methods.

Because of the high energy density and temperature that are required for the phase transition to a QGP, in most studies collisions of heavy-ions accelerated at ultra-relativistic energies are used to study the QGP. In a *Pb-Pb* collision at the Large Hadron Collider (LHC), the momentum of the two nuclei allows for a collision energy which is more than sufficient for such a plasma to form.

Recently however, there has been a sudden change of focus towards smaller systems, namely collisions between two protons: *pp* collisions. This change was triggered by experimental results that resemble signatures of a QGP in heavy-ion collisions. Precisely these same signatures are interpreted as the outcome of a QGP behaving like a liquid in heavy-ion collisions.

This is something which was unexpected, and thus very interesting. Because of the small size of the system in a *pp* collision, one was not expecting a QGP to form. Therefore it needs to be found out whether this actually is the case, or that the experimental signatures can be explained by conventional mechanisms already incorporated within QCD.

In order to get more insight into the answer to this question it is useful to look at the shape of *pp* collisions. There are two sorts of events: jetty and spherical. The jetty events are easily understood by QCD as a result of hard processes, and the spherical events can be described by phenomenological approaches. One of the mechanisms which could explain some of the plasma-like features found in *pp* collisions is called colour reconnection (CR).

This leads to the formulation of the following research question for this thesis:

"How can the shape observables of pp collisions generated by PYTHIA, a model often used to describe pp collisions, be used to discern between jetty and spherical events and thus describe the physical characteristics of a collision? Moreover, how does colour reconnection impact these observables?"

In this thesis, a description of the theory underlying *pp* collisions with relevant information is presented in section 2. Then, in section 3 the model used to generate the collision data is described along with the choices made for the analysis. In section 4 the results that were found for each observable are presented in histograms for each p_T and η selection. Lastly, in section 5 and section 6 it is concluded how the transverse sphericity can be used to distinguish between different types of collisions and the connection to colour reconnection is made.

2 Theory

2.1 Standard Model

The Standard Model of particle physics (SM) is the most fundamental and thorough description of the world around us so far. Except for a few notable exceptions it can be used to describe the Universe on the smallest scales and derive properties on larger scales [1]. Already mentioned, the exceptions that prevent this theory from being a 'theory of everything' are: gravity, the imbalance between matter and antimatter, neutrino oscillations and dark energy.

In brief, the theory describes all the matter and particles that are known to exist as fermions with a half-integer spin. How these particles interact is described as force carrying particles called bosons that have integer spins. The fermions are subdivided in quarks, the particles that make up atomic nuclei, and leptons consisting of the electron and its cousins muon and tau and their neutrino counterparts.

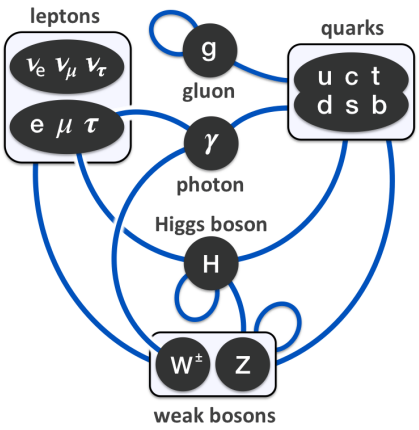


Figure 1 – Schematically representation of the Standard Model of particle physics. The ovals represent (groups of) particles, and the vertices possible interactions [2].

Of the four fundamental forces of nature three are included in the SM and each force is carried by a corresponding boson: the strong force by gluons, the weak force by $W^{\pm}$ and Z bosons and the electromagnetic force by photons. Gravity is the force that is not included here.

This system of particles and their relations is summarised in Figure 1. For every interaction-type there is a theory that describes how it works in detail. There is quantum chromodynamics (QCD) for the strong force, quantum electrodynamics (QED) for the electromagnetic force and Glashow-Weinberg-Salam theory for the weak force. Of these theories QCD is relevant for this research.

2.2 Quantum chromodynamics

The theory of QCD describes the strong interaction that governs the behaviour of quarks and gluons. It operates on a playing field consisting of six quark flavours (up, down, charm, strange, top, bottom). Similar to the more familiar electrodynamics, these quarks possess a unit of charge. In QCD however, the charge is called colour charge, of which there are three kinds instead of just one: red, blue and green. Each colour can take on a positive or a negative value (an anti-colour), but every quark only has one colour-type.

Colour can be transferred between quarks by gluons. Here they serve the same role as photons in electrodynamic interactions. However, in contrary to photons, gluon do carry charge themselves. Colour charge that is, and they carry both a positive and a negative unit of either kind. This means that gluons can interact with each other.

Charges of colour interact with each other with a strength corresponding to the coupling constant α_S. In Figure 2 the magnitude of this constant is shown for a range of momentum transfer-values, a variable that is inverse related to interaction length. Two features of this coupling constant give rise to interesting phenomena.

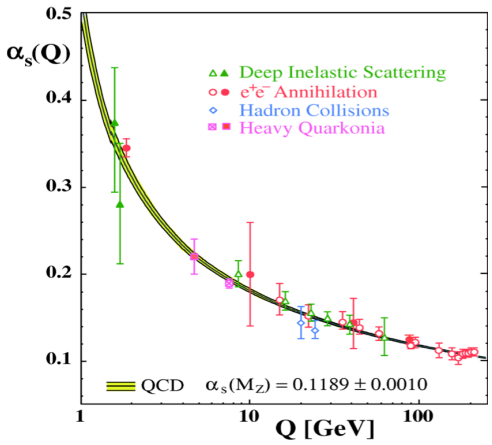


Figure 2 – Momentum transfer Q versus the QCD coupling constant α_S, where the line is the shape predicted by theory and the points are results from experiments. [3]

At low Q-values, and therefore high interaction lengths or distances, the strength of the strong nuclear force becomes incredibly high. As a result of this, separating two quarks from each other requires a large amount of energy. While pulling two quarks apart, this energy at some point becomes large enough for a quark anti-quark-pair to pop into existence in-between. Because of this, quarks are not allowed to move freely under normal circumstances and are bound into multiples called hadrons. This is called colour confinement.

At high Q-values, and therefore low interaction lengths or distances, the strength of the force becomes negligible. Because of the low value of α_S here perturbative QCD techniques can be used to theoretically predict the behaviour of quarks and gluons. On this scale the quarks do not interact significantly. This is called asymptotic freedom.

2.3 Quark-gluon plasma

Where normally quarks are governed by confinement, under extreme conditions de-confinement can occur, allowing asymptotic freedom to take over. Above a temperature of about 170 MeV and a energy density of 1 GeV/fm^3, normal matter transitions into so a quark-gluon plasma (QGP). In this phase, quark matter behaves like a perfect liquid. It is believed that the whole Universe was a QGP during the first fractions of a second after the Big Bang.

A QGP can be experimentally created by colliding two heavy-ions accelerated near the speed of light. At these ultra-relativistic speeds a collision will produce a 'fireball' of QGP, which will expand due to its initial pressure, and thus cool down after which confinement takes over control again causing hadronisation. The produced hadrons and other particles can then be detected.

2.4 Experiments

Experiments where QCD is probed are conducted all over the world. The largest collection of such experiments is situated in Geneva at the CERN laboratory. Here, scientists from all over the world participate in an European collaboration of nuclear research.

The main site of the experiments is the Large Hadron Collider (LHC). At a depth of 175 m, in a circular tunnel with a circumference of 27 km, the largest particle accelerator lies. Here particles can be collided

with a total energy of 13 TeV.

Within the LHC there are four main experiments: ATLAS, LHCb, CMS and ALICE. ATLAS and CMS were involved in the discovery of the Higgs boson, and are general-purpose particle detectors. LHCb looks at solving one of the mysteries that is not yet explained by the SM: the matter-antimatter asymmetry.

ALICE is optimised to look at the conditions where a QGP can be formed, so its properties such as deconfinement can be analysed. It has been steadily generating results from *Pb-Pb* collisions and currently also from *p-Pb* collisions and *pp* collisions. Within ALICE, 18 detectors are used to identify the produced particles, reconstruct their trajectories, and measure their energy [4].

2.5 Transverse sphericity

One of the observables of an event that closely describes the event shape of a collision is the transverse sphericity. Hence, it is a measure of the shape of the 'cloud' of the produced particles in a collision. In order to calculate the transverse sphericity, the transverse momentum matrix is constructed as follows [5]:

$$\mathbf{S_{xy}^{L}} = \frac{1}{\sum_i p_{Ti}} \sum_i \frac{1}{p_{Ti}} \begin{pmatrix} p_{xi}^2 & p_{xi}p_{yi} \\ p_{yi}p_{xi} & p_{yi}^2 \end{pmatrix}, \tag{1}$$

where p_{Ti} is the transverse momentum as defined earlier of a particle, p_{xi} and p_{yi} are the x- and y-components of the momentum respectively of a particle. This matrix is linearised in order to avoid a reduction in total momentum when two particles move in exactly the opposite direction.

Using the eigenvalues of this momentum matrix the transverse sphericity is calculated as follows:

$$S_T \equiv \frac{2\lambda_2}{\lambda_2 + \lambda_1}, \tag{2}$$

where $\lambda_1 > \lambda_2$.

The value of S_T ranges from 0 to 1, and these extremes correspond to two clear physical situations. When S_T is close to 0, the event is 'pencil'-like and will be referred to from now on as jetty. On the other hand, when S_T is close to 1, the event has an isotropic production of particles and will be referred to as spherical in the further text. In order to illustrate this, in Figure 3 several events are displayed for a range of sphericity values.

These events are plotted in $\eta - \phi$-space, where η is the pseudorapidity and ϕ is the azimuthal angle. This is a commonly used coordinate system in particle physics. The positions of the produced particles in a collision at a detector at the LHC can be represented in a Cartesian coordinate system where the z-axis is fixed along the beam direction that the particles are travelling along. From this system, the pseudorapidity can be calculated using the following equation:

$$\eta \equiv -\ln\left(\tan\left(\frac{\theta}{2}\right)\right), \tag{3}$$

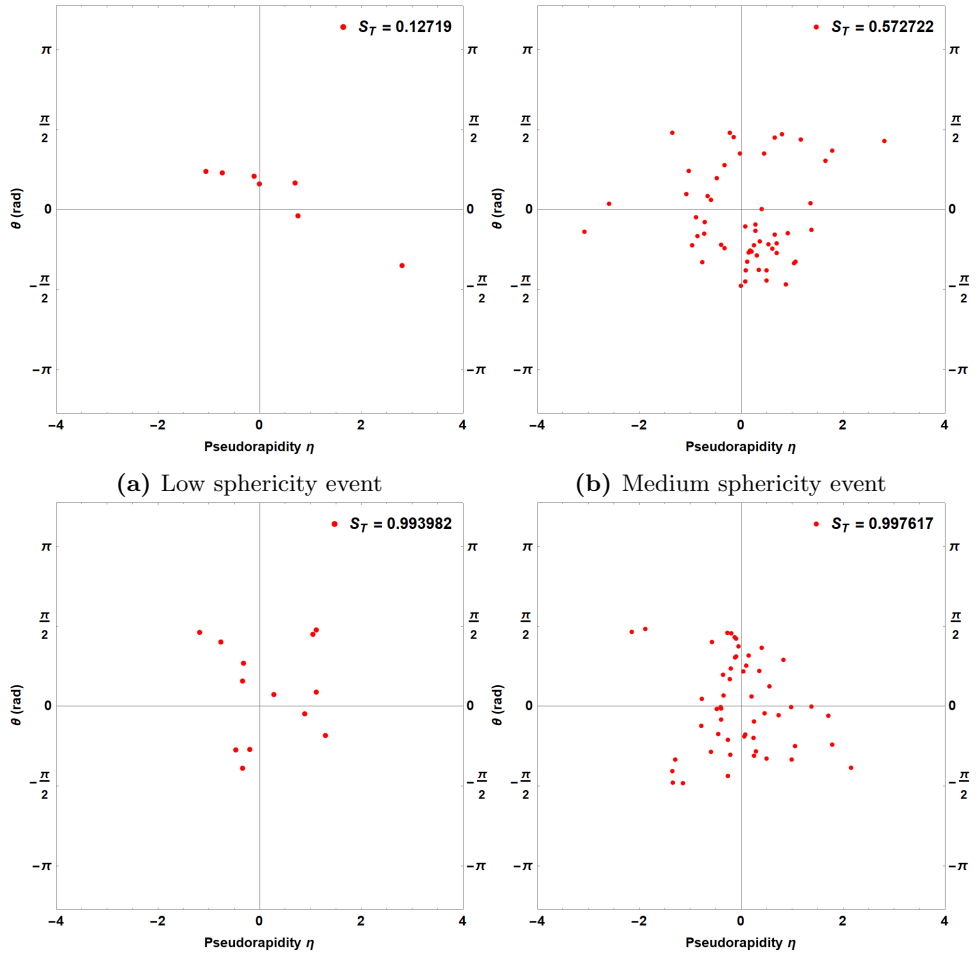
where the angle

$$\theta = \arccos\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right) \tag{4}$$

is used. The azimuthal angle is calculated as follows:

$$\phi = \arctan\left(\frac{y}{x}\right). \tag{5}$$

In these equations, x, y and z are the position coordinates of a particle, but the definition also allows these to be replaced with the momentum components of a particle.



(a) Low sphericity event **(b)** Medium sphericity event

(c) High sphericity event, low multiplicity **(d)** High sphericity event, high multiplicity

Figure 3 – Plots in $\eta - \phi$-space of particles produced in events with low, medium and high sphericity [6].

These events are generated with a model that produces a random set of Gaussian distributed momenta, resulting in a random set of produced particles. This process is repeated for a large amount of such events. For each event the sphericity is calculated. Even though there are no actual physics behind this model, it does provide a clear view of how events with different transverse sphericity values look like.

In the figure on the top left, a jetty event is displayed ($S_T = 0.13$), where the particles are not distributed uniformly. No particles occupy the third quadrant, yet many occupy the rest of the quadrants. The uneven distribution results in the low sphericity. In the top right of the figure a medium sphericity event is shown. The distribution is more uniform than in the previous one, however there is a large concentration of particles in the bottom right of the origin, indicating jetty behaviour.

In the bottom figures the distribution is balanced and uniform. Two highly events are shown, one with low multiplicity one with high multiplicity, in order to illustrate the multiplicity dependency of sphericity. Low sphericity events mainly occur for low multiplicities; high sphericity events can occur at either multiplicity.

3 Analysis

3.1 Material

For the analysis a series of observables was calculated on samples of Monte Carlo generated events. The software used to generate these events is PYTHIA 8.2, a dedicated program based on C++ for generating high-energy collisions between particles [7, 8].

Using PYTHIA, two sets of data have been generated, consisting of a lot of simulated pp collisions at $\sqrt{s} = 7$ TeV. Each dataset consists of a large amount (order of 10^8) of events (collisions between protons), which in turn consists of about $1-400$ primary produced particles. Each particle has a three-dimensional momentum as well as other properties which are irrelevant for this research.

In PYTHIA, the phenomenological mechanism of colour reconnection describes what happens to a set of quarks connected by colour-gluon-lines when one of the quarks flies away. In the case of colour reconnection, the quarks would reconnect to their nearest neighbours, allowing the 'leaving' quark to 'drag along' nearby quarks. When colour reconnection is not incorporated into the model, the lines just disappear.

One of the datasets has colour reconnection turned on, here around 4.0×10^8 events were generated. In the other dataset, colour reconnection was turned off, here 6.7×10^8 events were generated.

The datasets were analysed using ROOT, the software package of CERN that allows for fast processing of large amounts of data and which is also based on C++ [9].

3.2 Method

For each measurement, a different selection on the available phase space is made. These selections are applied to the transverse momentum and pseudorapidity.

For the transverse momentum the following ranges are used as intervals: $2 \leq p_T \leq 10$ (high p_T), $0 \leq p_T \leq 2$ (low p_T) and $0 \leq p_T \leq 10$ (full-range p_T), all in GeV/c. The low p_T range is used because the bulk of the produced particles is concentrated in this range. Because of the high amount of particles it is possible to study this range in detail with a Monte Carlo-generated model. In the high p_T range the particles can be described using perturbative QCD techniques.

Moreover, three different pseudorapidity ranges are used, that can be seen in Figure 4. These ranges correspond to the acceptance ranges of two detectors in the LHC: $|\eta| \leq 0.8$ is the acceptance of ALICE and $|\eta| \leq 2.4$ is the acceptance of CMS. The range of $|\eta| \leq 10$ corresponds to the entire phase space.

Table 1 – Momentum and pseudorapidity selections used in the analysis. The ordering is the same as in the upcoming figures.

Cut №	$p_{T,min}$	$p_{T,max}$	η_{min}	η_{max}
1	2	10	-0.8	0.8
2	2	10	-2.4	2.4
3	2	10	-10	10
4	0	2	-0.8	0.8
5	0	2	-2.4	2.4
6	0	10	-10	10

The different selections are shown in Table 1, where each cut corresponds to a different run of the data analysis software.

The data is processed according to the following procedure. For each particle in an event, the transverse momentum is calculated using the following formula:

$$p_T = \sqrt{p_x^2 + p_y^2}, \tag{6}$$

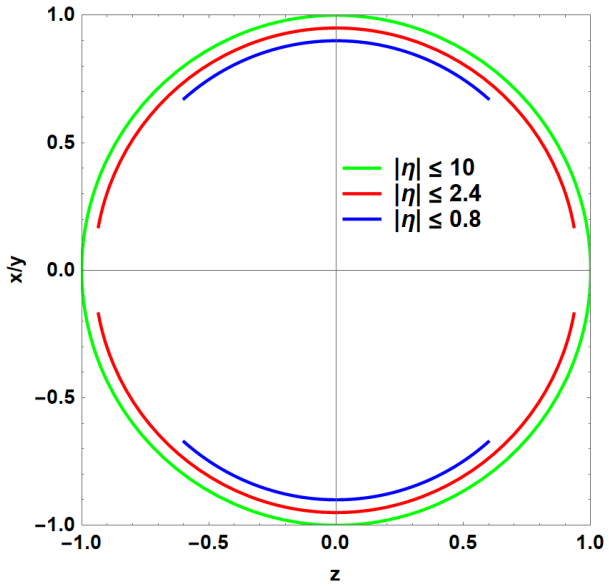


Figure 4 – Visual representation of the pseudorapidity cuts in the z-x/y-plane. Blue is for pseudorapidity $-0.8 \leq \eta \leq 0.8$ as in Alice, red is for $-2.4 \leq \eta \leq 2.4$ as in CMS and green is for $-10.0 \leq \eta \leq 10.0$ which takes 99.99% of the full angular range.

where p_x and p_y are the x- and y-components of the momentum.

Then, using the calculation method described in subsection 2.5, for each event the sphericity is calculated. The mean p_T is also computed for each event. These two observables, along with the multiplicity (the amount of primary produced particles in an event), are then counted in a set of histograms.

4 Results

In the following section the results of the analysis are presented. All results in this section used the PYTHIA generator with colour reconnection turned on. In the appendix the results with colour reconnection turned off are shown, however these do not differ a lot from the results presented here.

In Figure 5 a response matrix is displayed that shows the relation between the theoretical, full sphericity and the measured sphericity. The measured sphericity is calculated by dropping two out of ten particles from an event randomly. This corresponds to an efficiency of 80%, which is similar to that of most detectors at the LHC. The line $y = x$ follows the distribution of a histogram of the sphericity. The anomalies appear in the off-diagonal entries. There are more of these anomalies above the line, where the theoretical sphericity is larger than the measured sphericity. Note that this sphericity is not what is measured in this study, but what a detector would measure with actual particle collisions. Thus this response matrix can be used to correct the measured values of S_T and allow a direct comparison with theoretical calculations.

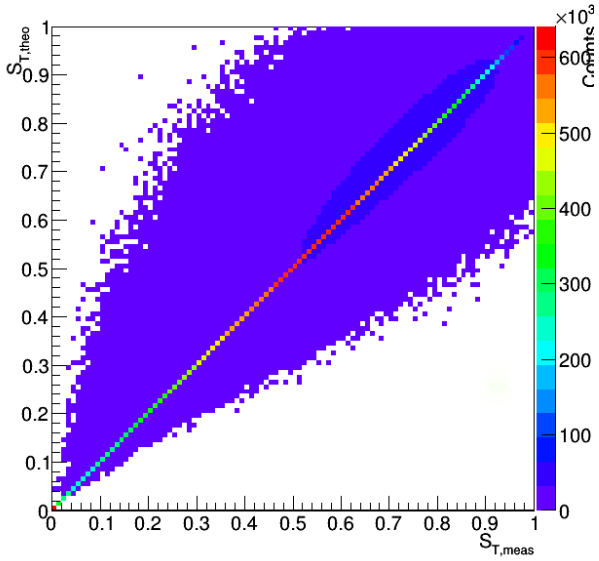


Figure 5 – Response matrix at 80% efficiency. The used cuts are $0 \leq p_T \leq 2$ and $-0.8 \leq \eta \leq 0.8$. Colour reconnection is turned on.

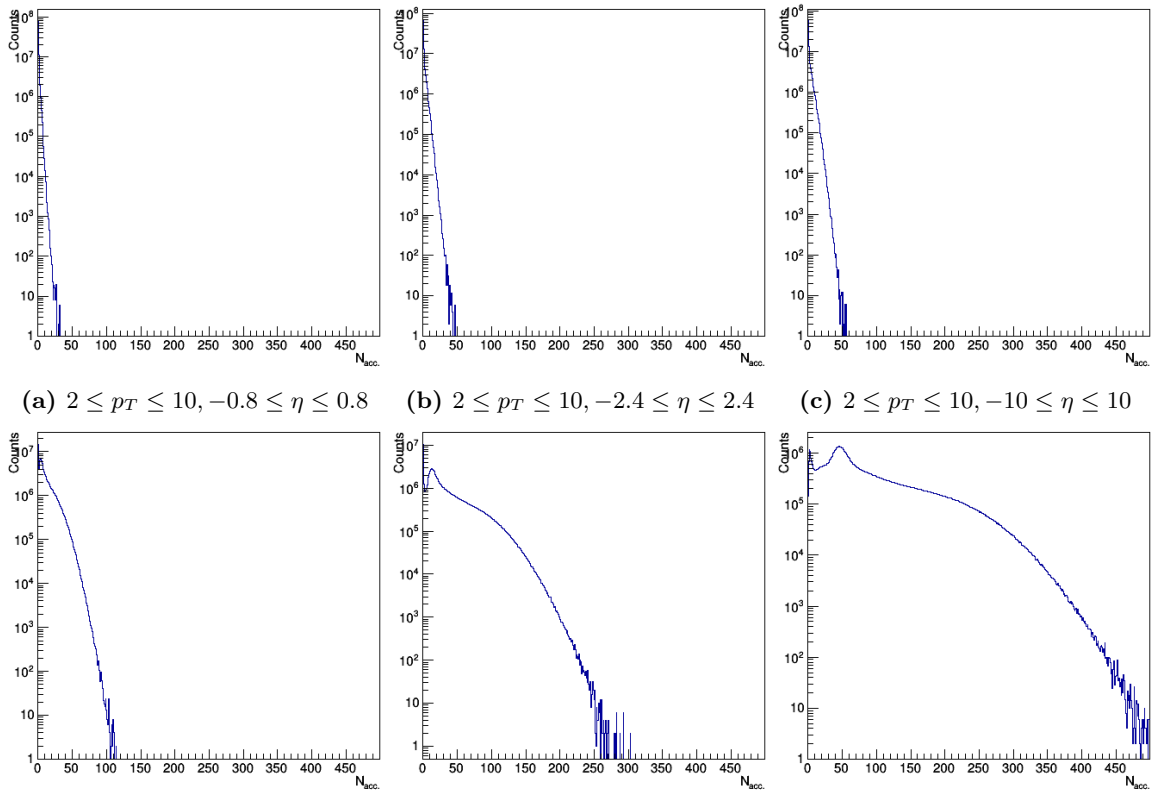
4.1 Multiplicity

In Figure 6 we see the distribution of different multiplicities of the events. For the high p_T histograms it can be seen that most events have a low multiplicity. The maximum multiplicity increases slightly if the pseudorapidity, the angular range is increased, from 35 events to 55 events.

The case is much different when a low p_T cut is applied. High multiplicities now come available. Moreover, as the angular range is amplified, the maximum multiplicity increases a lot. An interesting observation that becomes clearer at higher pseudorapidity values is that there is a second peak of multiplicity count with about the same height as the primary peak. The primary peak occurs at 1/2/3 multiplicity, and the secondary appears one order of magnitude further away.

At the higher multiplicities a higher noise of the line is observed. This is logical since here the amount of events is lower, an effect which is amplified by the log-y-axis.

The amount of primary produced particles clearly depends on the transverse momentum of these particles. There are no events with high p_T that have a high multiplicity.



(a) $2 \leq p_T \leq 10, -0.8 \leq \eta \leq 0.8$ **(b)** $2 \leq p_T \leq 10, -2.4 \leq \eta \leq 2.4$ **(c)** $2 \leq p_T \leq 10, -10 \leq \eta \leq 10$

(d) $0 \leq p_T \leq 2, -0.8 \leq \eta \leq 0.8$ **(e)** $0 \leq p_T \leq 2, -2.4 \leq \eta \leq 2.4$ **(f)** $0 \leq p_T \leq 10, -10 \leq \eta \leq 10$

Figure 6 – Multiplicity histograms for each selection. The 'counts'-axis is logarithmic.

4.2 Sphericity

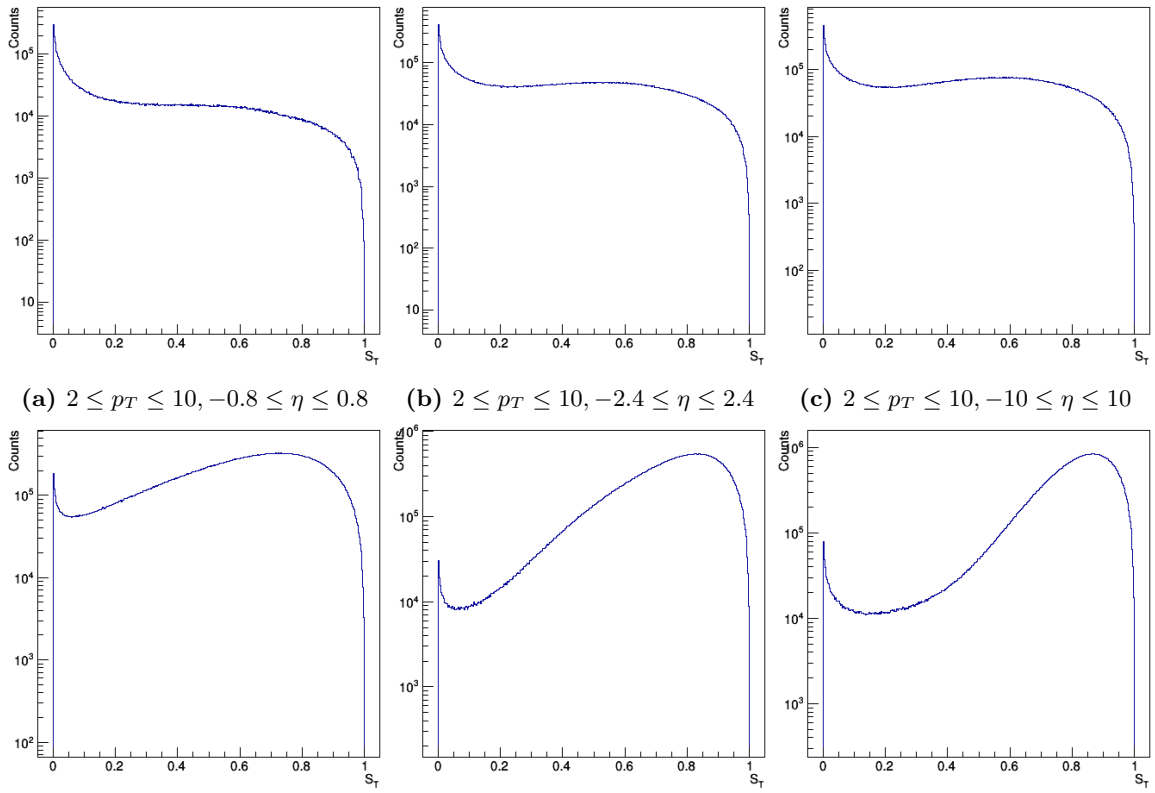
In Figure 7 the distribution of the different sphericity values can be seen. There is a clear difference in the shape of the curve between high-p_T and low-p_T cuts. This difference is that while at the high p_T cut the sphericity has a maximum at a very low value, whereas in the low p_T cut there is a second maximum near the highest sphericity values. The overall shape is the same across pseudorapidity ranges. However, the shape of this second peak becomes more pronounced at higher pseudorapidity ranges.

This difference between the high and low transverse momentum ranges is related to the difference in the multiplicity histograms. Apparently high sphericity events occur more at high multiplicity events, which occur more at the low p_T range.

4.3 Multiplicity versus sphericity

In Figure 8 the 2D histograms of multiplicity versus sphericity are shown. For the high p_T range the values are uniformly distributed along the low multiplicity values across the whole spectrum of sphericity values. This is logical, since in the histogram of multiplicity it was seen that no high multiplicity events occur at high values for p_T. However, it is interesting that even at these low multiplicity events high sphericity is also observed.

At the low p_T cut a histogram is produced with more characteristics. Spherical events not necessarily happen at high multiplicity values, however the reverse is the case. High multiplicity events clearly are spherical. This is logical since the chance for a large amount of emitted particles to be emitted in the same direction is low. It can also be seen that the highest multiplicity events are only observed in the highest pseudorapidity range.



(a) $2 \leq p_T \leq 10, -0.8 \leq \eta \leq 0.8$ **(b)** $2 \leq p_T \leq 10, -2.4 \leq \eta \leq 2.4$ **(c)** $2 \leq p_T \leq 10, -10 \leq \eta \leq 10$

(d) $0 \leq p_T \leq 2, -0.8 \leq \eta \leq 0.8$ **(e)** $0 \leq p_T \leq 2, -2.4 \leq \eta \leq 2.4$ **(f)** $0 \leq p_T \leq 10, -10 \leq \eta \leq 10$

Figure 7 – Transverse sphericity histograms for each selection. The 'counts'-axis is logarithmic. S_T is calculated as in subsection 2.5.

4.4 Mean transverse momentum versus sphericity

In Figure 9 the 2D histograms of the transverse momentum versus sphericity are displayed. As can be seen in Figure 9f, where the distribution is seen across the whole range of values, where no cut is applied, most events are observed in the low p_T range. In the case of high p_T, the majority of the events has a low sphericity and is among the low-end of this momentum-range. This can be seen by a clear peak. At the highest pseudorapidity range a second peak at mid-high sphericity is observed.

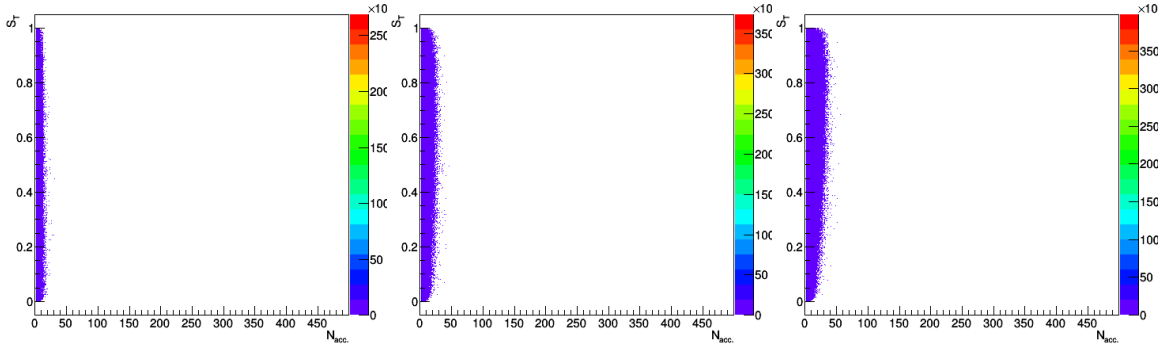
At the low p_T range a clear peak is observed around $(p_T, S_T) = (0.4, 0.9)$. This peak becomes more pronounced at broader pseudorapidity ranges.

The position of the peak indicates that high p_T means low sphericity, thus jetty events. This means that high p_T events, mainly originate from a hard scattering events and thus have a higher likelihood of producing jets.

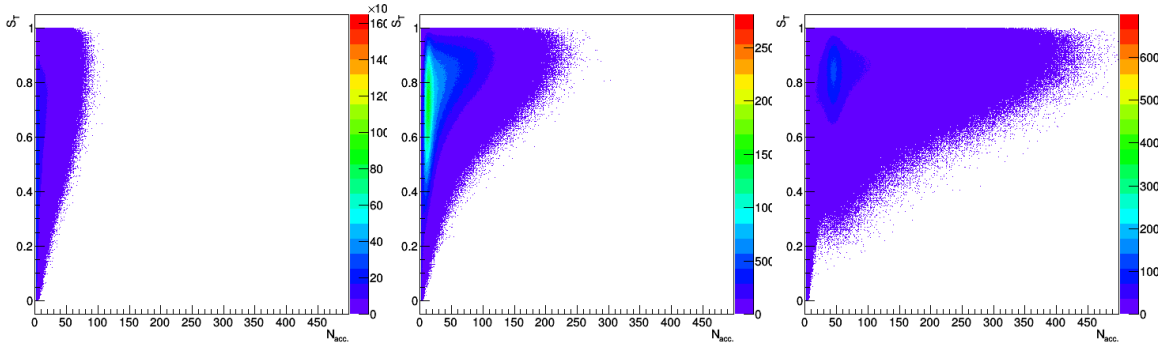
4.5 Comparison between colour reconnection on and off

In Figure 10 the mean values of p_T, S_T, and multiplicity are displayed for both high and low momentum ranges, and for both CR and No CR. It can be seen that all differences between CR and No CR are the other way around for each momentum range.

The largest difference is seen in the low p_T range. Here, the mean sphericity and mean multiplicity are much smaller when CR is turned on.

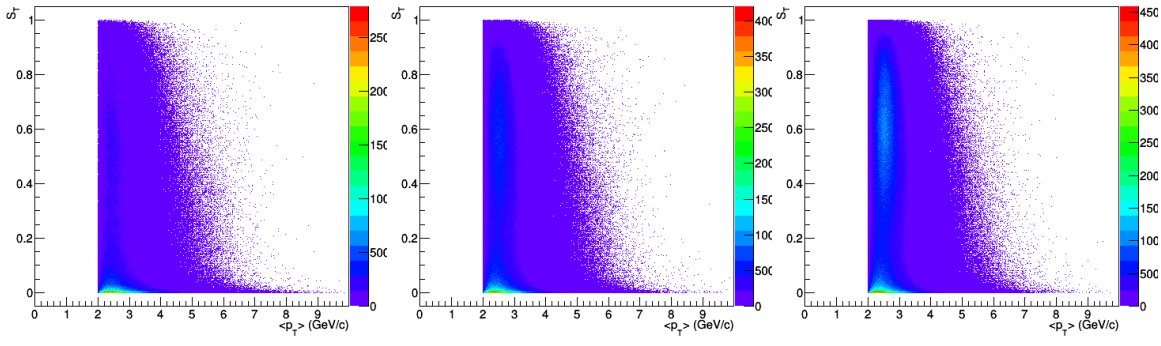


(a) $2 \leq p_T \leq 10, -0.8 \leq \eta \leq 0.8$ **(b)** $2 \leq p_T \leq 10, -2.4 \leq \eta \leq 2.4$ **(c)** $2 \leq p_T \leq 10, -10 \leq \eta \leq 10$

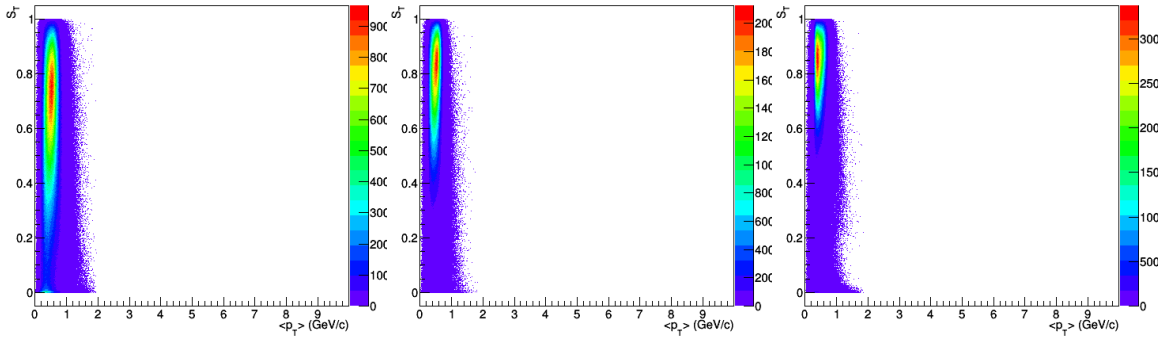


(d) $0 \leq p_T \leq 2, -0.8 \leq \eta \leq 0.8$ **(e)** $0 \leq p_T \leq 2, -2.4 \leq \eta \leq 2.4$ **(f)** $0 \leq p_T \leq 10, -10 \leq \eta \leq 10$

Figure 8 – Multiplicity versus sphericity 2D-histograms for each selection.

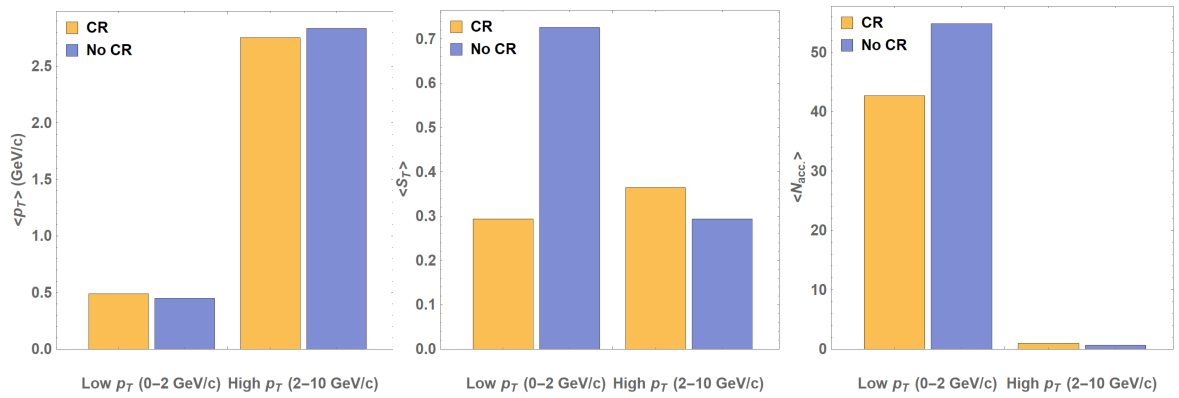


(a) $2 \leq p_T \leq 10, -0.8 \leq \eta \leq 0.8$ **(b)** $2 \leq p_T \leq 10, -2.4 \leq \eta \leq 2.4$ **(c)** $2 \leq p_T \leq 10, -10 \leq \eta \leq 10$



(d) $0 \leq p_T \leq 2, -0.8 \leq \eta \leq 0.8$ **(e)** $0 \leq p_T \leq 2, -2.4 \leq \eta \leq 2.4$ **(f)** $0 \leq p_T \leq 10, -10 \leq \eta \leq 10$

Figure 9 – Mean transverse momentum versus sphericity 2D-histograms for each selection.



(a) Mean transverse momentum **(b)** Mean transverse sphericity **(c)** Mean multiplicity

Figure 10 – Bar charts of respectively the mean transverse momentum, sphericity and multiplicity for both a the high and low transverse momentum selection. These values are included for the results of the model with CR turned on and CR turned off (No CR).

5 Conclusion

Out of the various observables of a *pp* collision, a lot of insight in the characteristics of this collision can be found. Using the sphericity, the distinction can be made between spherical events and jetty events. This distinction is clearly related to the other observables. Low momentum events are mostly spherical, and high momentum events are mostly jetty. Moreover, high multiplicity events are preferentially spherical.

6 Discussion and outlook

For the qualitative analysis that is mostly done in this research, errors do not play a large role. Moreover, due to the large amount of events in each dataset, the results are accurate enough to draw conclusions from.

In order to be able to better identify the difference that colour reconnection makes, the relationship between mean p_T and multiplicity needs to be researched. In Figure 11 such a plot is shown. It can be seen that only when CR is turned on, the mean p_T rises with multiplicity. This is an effect which explains some of the flow-like behaviour in *pp* collisions and can be attributed to CR [10, 11, 12].

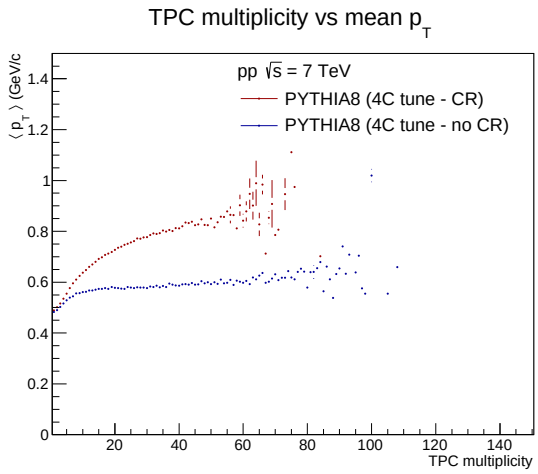


Figure 11 – Multiplicity versus mean transverse momentum plot for PYTHIA 8 generated data with CR turned on (red line) and CR turned off (blue line) [10]

Moreover, in further research the multiplicity histograms could be normalised in order to be able to compare the CR data with the no CR data. This could be interesting because in the not normalised histograms it seems that when CR is turned off, the maximum multiplicity is higher.

References

[1] E. M. Henley and A. Garcia. *Subatomic physics*. World Scientific, 2007.

[2] E. Drexler. *Elementary particle interactions in the Standard Model*. Wikimedia Commons. 17th Apr. 2014. URL: https://commons.wikimedia.org/wiki/File:Elementary_particle_interactions_in_the_Standard_Model.png.

[3] G. M. Prosperi, M. Raciti and C. Simolo. 'On the running coupling constant in QCD'. In: *Prog. Part. Nucl. Phys.* 58 (2007), pp. 387–438. DOI: 10.1016/j.ppnp.2006.09.001. arXiv: hep-ph/0607209 [hep-ph].

[4] ALICE Collaboration et al. 'ALICE: Physics Performance Report, Volume I'. In: *Journal of Physics G: Nuclear and Particle Physics* 30.11 (2004), p. 1517. URL: http://stacks.iop.org/0954-3899/30/i=11/a=001.

[5] B. Abelev et al. 'Transverse sphericity of primary charged particles in minimum bias proton–proton collisions at $\sqrt{\mathbf{s}} = \mathbf{0.9}$, **2.76** and **7** TeV'. In: *The European Physical Journal C* 72.9 (2012), p. 2124. URL: https://link.springer.com/article/10.1140/epjc/s10052-012-2124-9.

[6] C.-Y. Wong. *Introduction to high-energy heavy-ion collisions. Pseudorapidity Variable*. World Scientific, 1994. Chap. 2.4, pp. 24–26.

[7] T. Sjöstrand. *PYTHIA 8. Welcome to PYTHIA - The Lund Monte Carlo!* Version 8.186. Lund University. 2014. URL: http://home.thep.lu.se/~torbjorn/pythia81html/Welcome.html.

[8] T. Sjöstrand, S. Mrenna and P. Z. Skands. 'A Brief Introduction to PYTHIA 8.1'. In: *Comput. Phys. Commun.* 178 (2008), pp. 852–867. DOI: 10.1016/j.cpc.2008.01.036. arXiv: 0710.3820 [hep-ph].

[9] R. Brun and F. Rademakers. *ROOT Users Guide. Data Analysis Framework*. CERN, Geneva. 2007. URL: https://root.cern.ch/root/htmldoc/guides/users-guide/ROOTUsersGuide.html.

[10] A. Ortiz Velasquez et al. 'Color Reconnection and Flowlike Patterns in *pp* Collisions'. In: *Phys. Rev. Lett.* 111.4 (2013), p. 042001. DOI: 10.1103/PhysRevLett.111.042001. arXiv: 1303.6326 [hep-ph].

[11] T. Sjöstrand. 'Colour reconnection and its effects on precise measurements at the LHC'. In: 2013. arXiv: 1310.8073 [hep-ph]. URL: https://inspirehep.net/record/1262718/files/arXiv:1310.8073.pdf.

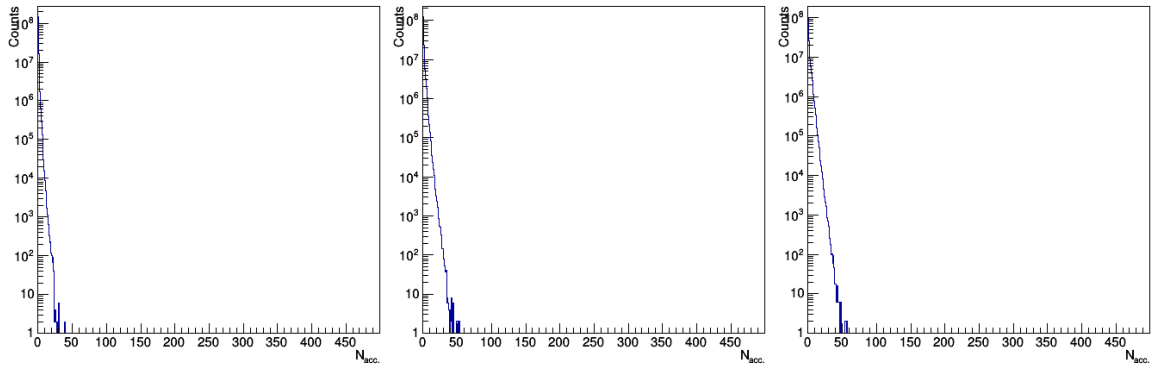
[12] E. Cuautle et al. 'Color reconnection: a fundamental ingredient of the hadronisation in p-p collisions'. In: *Journal of Physics: Conference Series* 730.1 (2016), p. 012009. URL: http://stacks.iop.org/1742-6596/730/i=1/a=012009.

A Colour reconnection off results

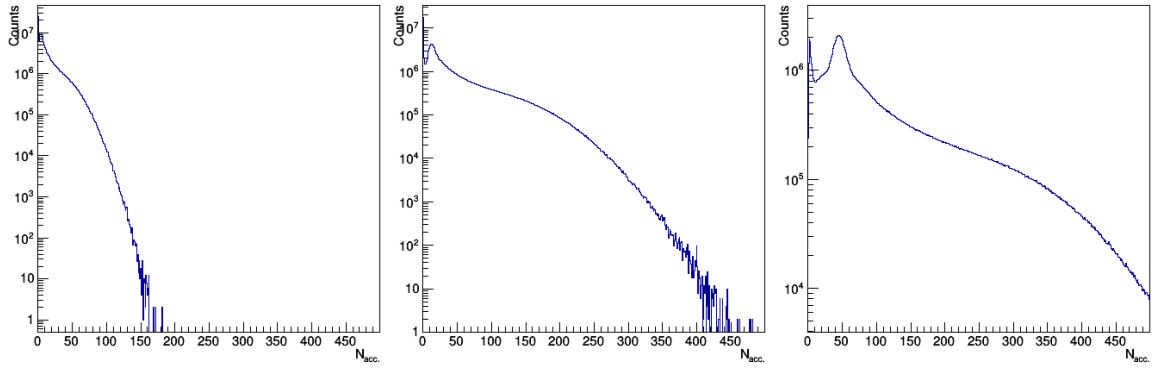
In Figure 12 until Figure 15 the results are presented where PYTHIA with colour reconnection turned off was used to generate the events. The shape of the graphs is mostly the same as with CR turned on.

A.1 Multiplicity

In Figure 12 the multiplicity histograms are displayed.



(a) $2 \leq p_T \leq 10, -0.8 \leq \eta \leq 0.8$ **(b)** $2 \leq p_T \leq 10, -2.4 \leq \eta \leq 2.4$ **(c)** $2 \leq p_T \leq 10, -10 \leq \eta \leq 10$



(d) $0 \leq p_T \leq 2, -0.8 \leq \eta \leq 0.8$ **(e)** $0 \leq p_T \leq 2, -2.4 \leq \eta \leq 2.4$ **(f)** $0 \leq p_T \leq 10, -10 \leq \eta \leq 10$

Figure 12 – Multiplicity histograms for each selection with CR turned off. The 'counts'-axis is logarithmic.

A.2 Sphericity

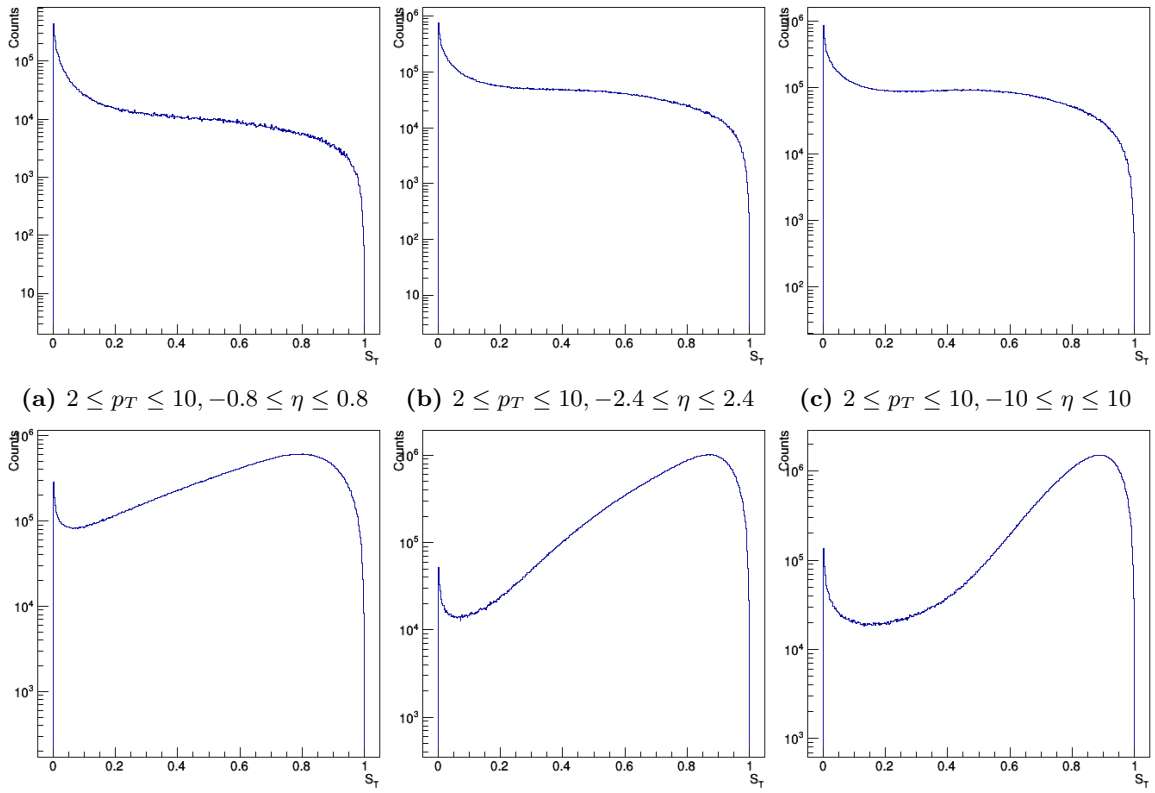
In Figure 13 the histograms for the transverse sphericity are displayed.

A.3 Multiplicity versus sphericity

In Figure 14 the 2D histograms of the multiplicity versus the transverse sphericity are displayed.

A.4 Mean transverse momentum versus sphericity

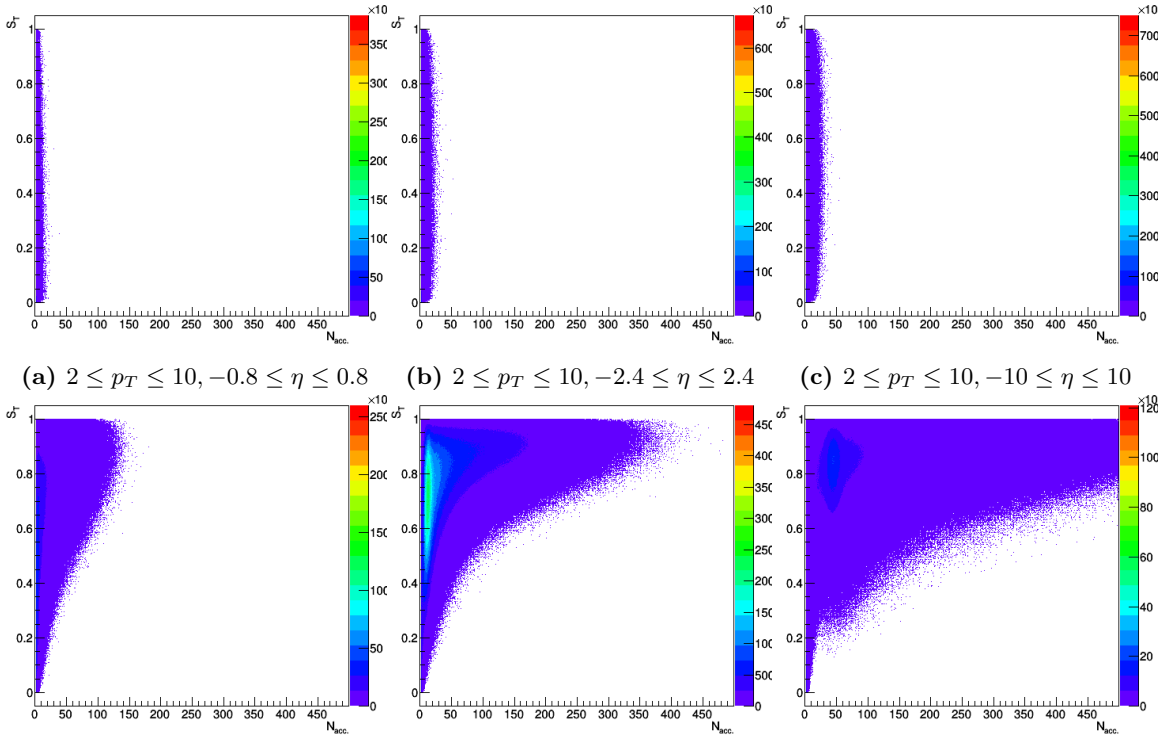
In Figure 15 the 2D histograms of the transverse momentum versus the transverse sphericity are displayed.



(a) $2 \leq p_T \leq 10, -0.8 \leq \eta \leq 0.8$ **(b)** $2 \leq p_T \leq 10, -2.4 \leq \eta \leq 2.4$ **(c)** $2 \leq p_T \leq 10, -10 \leq \eta \leq 10$

(d) $0 \leq p_T \leq 2, -0.8 \leq \eta \leq 0.8$ **(e)** $0 \leq p_T \leq 2, -2.4 \leq \eta \leq 2.4$ **(f)** $0 \leq p_T \leq 10, -10 \leq \eta \leq 10$

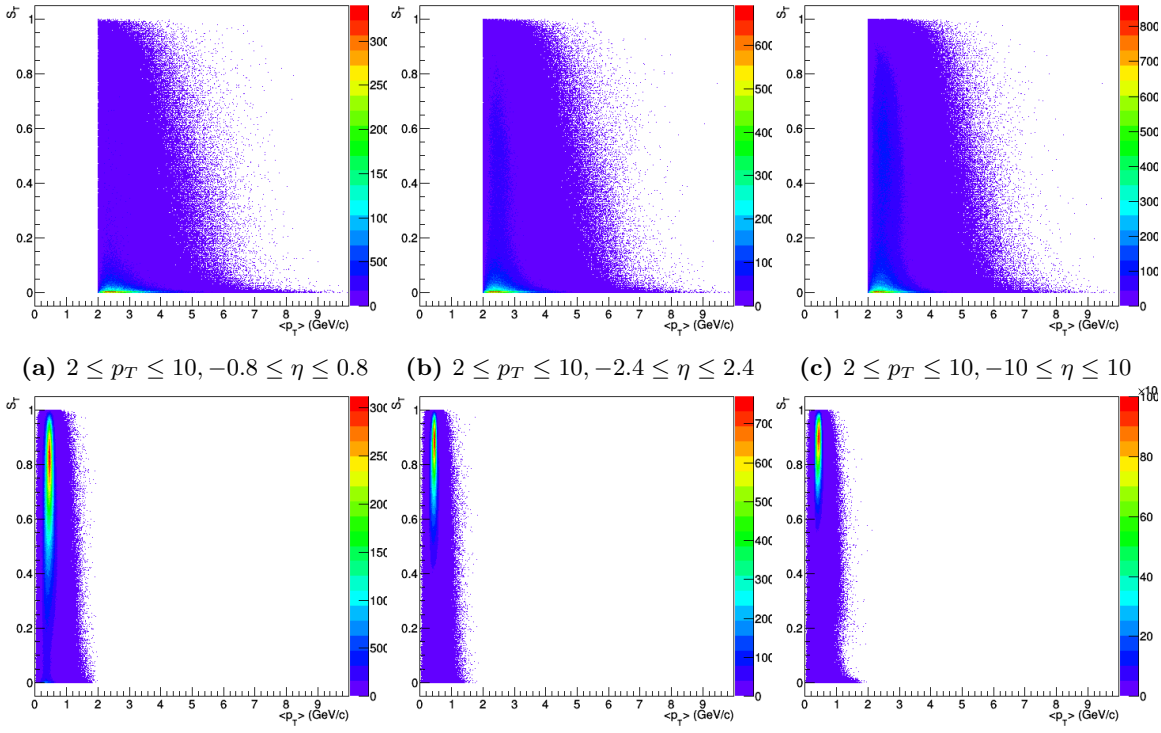
Figure 13 – Sphericity histograms for each selection with CR turned off. The 'counts'-axis is logarithmic.



(a) $2 \leq p_T \leq 10, -0.8 \leq \eta \leq 0.8$ **(b)** $2 \leq p_T \leq 10, -2.4 \leq \eta \leq 2.4$ **(c)** $2 \leq p_T \leq 10, -10 \leq \eta \leq 10$

(d) $0 \leq p_T \leq 2, -0.8 \leq \eta \leq 0.8$ **(e)** $0 \leq p_T \leq 2, -2.4 \leq \eta \leq 2.4$ **(f)** $0 \leq p_T \leq 10, -10 \leq \eta \leq 10$

Figure 14 – Multiplicity versus sphericity 2D-histograms for each selection with CR turned off.



(a) $2 \leq p_T \leq 10, -0.8 \leq \eta \leq 0.8$ **(b)** $2 \leq p_T \leq 10, -2.4 \leq \eta \leq 2.4$ **(c)** $2 \leq p_T \leq 10, -10 \leq \eta \leq 10$

(d) $0 \leq p_T \leq 2, -0.8 \leq \eta \leq 0.8$ **(e)** $0 \leq p_T \leq 2, -2.4 \leq \eta \leq 2.4$ **(f)** $0 \leq p_T \leq 10, -10 \leq \eta \leq 10$

Figure 15 – Mean transverse momentum versus sphericity 2D-histograms for each selection with CR turned off.